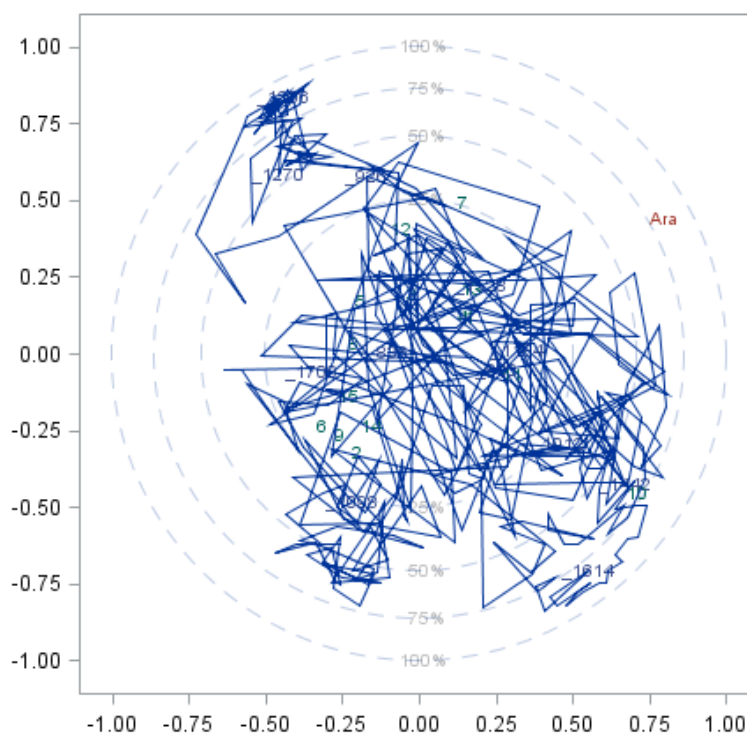


## FTIR- och Ramanspektroskopi som analysmetod för den kemiska sammansättningen av granved – en jämförelse

*FTIR- and Raman spectroscopy as analytical methods for the chemical composition of Norwegian spruce wood – a comparison*



Edward Carlsson & Douglas Kärrholmen

Sveriges Lantbruksuniversitet  
Institutionen för skogens ekologi och skötsel  
Kandidatarbete i skogsvetenskap, 15 hp,  
Handledare: Rosario Garcia Gil, SLU, Institutionen för skoglig genetik och växtfysiologi  
Bitr.handledare: Tomas Funda, SLU, Institutionen för skoglig genetik och växtfysiologi

Program: Jägmästarprogrammet

Kurs: EX0813 Nivå:G2E

Umeå 2018



# Kandidatarbeten i Skogsvetenskap

Fakulteten för skogsvetenskap,  
Sveriges lantbruksuniversitet

Examinator/Examiner	Institutionen för skogens ekologi och skötsel/ Department of Forest Ecology and Management
Författare/Author	Edward Carlsson & Douglas Kärrholmen
Titel, Sv	FTIR- och Ramanspektroskopi som analysmetod för den kemiska sammansättningen av granved – en jämförelse.
Titel, Eng	<i>FTIR- and Raman spectroscopy as analytical methods for the chemical composition of Norwegian spruce wood – a comparison</i>
Nyckelord	<i>Vibrationsspektroskopi, korsvalidering, vedkemi, absorbtion, närinfraröd, modell.</i>
Keywords	<i>Vibrational spectroscopy, cross validation, wood chemistry, absorbtion, near-infrared model.</i>
Handledare/Supervisor	Rosario Garcia Gil - Institutionen för skoglig genetik och växtfysiologi
Bitr handledare/Asst Supervisor	Tomas Funda - Institutionen för skoglig genetik och växtfysiologi.
Examinator/Examiner	Tommy Mörling Institutionen för skogens ekologi och skötsel/ Department of Forest Ecology and Management
Kurstitel/Course	Kandidatarbete i skogsvetenskap/ Bachelor Degree in Forest Science
Kurskod	EX0813
Program	Jägmästarprogrammet
Omfattning på arbetet	15 hp
Nivå och fördjupning på arbetet	G2E
Utgivningsort	Umeå
Utgivningsår	2018

## SAMMANFATTNING

Den kemiska sammansättningen har en central påverkan på egenskaper och användningsområden för trä. För att kunna uppfylla dagens krav för vedens kemiska sammansättning krävs billiga och effektiva analysmetoder. I denna studie jämfördes prediktionsmodeller för halten av arabinos, galaktos, glukos, mannos, cellulosa, hemicellulosa, lignin och extraktivämnen baserad på data genererat med två metoder av vibrationsspektroskopi, FTIR och Raman. Fem olika borkärnor av granar (*Picea abies*), uppdelade i tre regioner, användes som provmaterial. Från dessa skapades tre dataset, FTIR med malt träpulver mixat med kaliumbromid, Raman med rent träpulver och Raman med borkärnor. Modellerna utarbetades genom att skapa principiella faktorer i programvaran SAS och utvärderades baserat på deras  $R^2$  efter att PRESS minimerats. Majoriteten av alla modeller misslyckades med att ta fram ett värde för  $R^2$ . Av de modeller som skapades presterade FTIR bäst för mannos och lignin, och Raman för arabinos, glukos, xylos, extraktivämnen, cellulosa och hemicellulosa. Varken FTIR eller Raman kunde skapa en modell för galaktos. Ingen av metoderna framstod som entydigt bättre. Förbehandling av data samt val av provmaterial visade sig kunna ha en betydande inverkan på modellens styrka för specifika kemikalier. Bristerna i resultatet tillskrivs främst den begränsade provstorleken. Båda metoderna ansågs dock vara relevanta och intressanta för fortsatta studier, men endast med en större provstorlek.

Nyckelord: Vibrationsspektroskopi, korsvalidering, vedkemi, absorbtion, närinfraröd, modell.

## ABSTRACT

The chemical composition plays a key role for the properties and applications of wood. In order to meet contemporary demands on the chemical composition of wood, cheap and effective methods of analysis are required. In this study, we compared predictive models for arabinose, galactose, glucose, mannose, cellulose, hemicellulose, lignin and extractives content developed from data generated with two forms of vibrational spectroscopy, FTIR and Raman. Five different wood cores from Norway spruce trees, each one split into three regions, were used as test material. From these, three data sets were created; FTIR with wood powder mixed with sodium bromide, Raman with pure wood powder and Raman with wood cores. The models were developed by creating principal factors in the software SAS, and evaluated based on their  $R^2$  after PRESS was minimized. The majority of all developed models did not meet the requirements to calculate  $R^2$ . Of the models that were created, FTIR had the best performance for mannose and lignin, and Raman for arabinose, glucose, xylose, extractives, cellulose and hemicellulose. Neither FTIR nor Raman managed to create a model for galactose. None of the methods appeared unequivocally superior. The preprocessing of data and sample material of choice proved to greatly affect models for specific chemicals. Deficiencies in the results were mainly attributed to the limited sample size. However, both methods were considered relevant and interesting for continued studies, but only with larger sample sizes.

**Key words:** Vibrational spectroscopy, cross validation, wood chemistry, absorption, nearinfrared model.

# INTRODUKTION

Trä är en industriellt viktig råvara med många användningsområden. Den utnyttjas som grund till ett flertal produkter, främst som konstruktionsvirke och pappersmassa (Ellingwood et al. 2004) (Weinstock et al. 1997). Utöver dessa används trä även som en biobaserad energikälla, textilmaterial och för produktion av kemikalier (Baratieri et al. 2008) (Lupoi et al. 2015), (Binder 2009) (Rojo et al. 2013). Beroende på modifikation och första användningsområde kan träråvaran potentiellt återvinnas för flera syften under en livscykel, för att vid sitt slutskede förbrännas och bidra med energi (Ramage et al. 2017).

I Sverige är gran (*Picea abies*) ett av de mest utbredda och kommersiellt viktigaste trädslagen. År 2017 utgjordes det levande virkesförrådet i Sverige till 40.8 % gran (Fridman et al. 2017). Således är granen ett relevant trädslag för vidare forskning och utveckling.

## Uppbyggnad av trä

Det som i dagligt tal kallas för "trä" utgörs av sekundär xylemvävnad hos trädet. Vävnadens funktion är framförallt att transportera vatten från rötter till blad, där det kombineras med koldioxid och solenergi för att bygga sockermolekyler. För att utföra detta har cellerna i xylemvävnaden en tjock sekundär cellvägg för stadga, samt att ej ge vika för tryck vid transport av vatten. Denna utgörs främst av tre typer av kolbaserade makromolekyler; cellulosa, hemicellulosa och lignin. Den absoluta mängden av dessa tre molekyler varierar på art-, individ- och cellnivå. Vanligtvis består den sekundära cellväggen av 40-80 % cellulosa, 10-40 % hemicellulosa och 5-25 % lignin (Kumar et al. 2016). Cellulosa består av linjära kedjor av beta-1-4 glukos. Dessa binder sig samman till långa kristallina strukturer kallade mikrofibriller. Hemicellulosa tvärbinder samman mikrofibrillerna med vätebindningar. Strukturen hos hemicellulosa är en starkt diversifierad uppbyggnad av sockerkedjor av monosackariderna; xylos, glukos, mannos, galaktos, arabinos, och rhamnos. Hemicellulosa varierar mellan artgrupper, till exempel är den huvudsakliga formen av hemicellulosa galactoglucomannan i barrträ och arabinoglucuronoxylan i lövträ (Ramage et al. 2017). Lignin har en kraftigt amorf, heterogen och komplex kolstruktur vilken fyller ut hållrummen mellan mikrofibrillerna och hemicellulosan. Molekylen är svårnedbrytbar och gör cellväggen motståndskraftig mot mikrobiella angrepp (Lupoi et al. 2015). Resten av den sekundära cellväggen består främst små halter extraktivämnen, vilka inkluderar olika grupper av organiska molekyler som; färgämnen, fenoler, terpentiner, fettsyror, hartsämnen och metaboliska intermediärer (Ramage et al. 2017).

Den sekundära cellväggen spelar en central roll för träets materialegenskaper. Beroende på det vilka egenskaper som efterfrågas kan olika kompositioner av den sekundära cellväggen vara önskvärda. Exempelvis efterfrågas en låg ligninhalt för trä som används inom produktion av pappersmassa, eller för utvinning av biobränsle till kraftvärmeverk (Weinstock et al. 1997) (Gierlinger 2014).

Sammansättningen av trä kan idag bestämmas med god noggrannhet med hjälp av olika former av våtkemiska analyser. Problemet är att metoden kan vara kostsam, tidskrävande och komplicerad (Hames et al. 2003). Processen riskerar även att orsaka skador på trämassan och att skapa restprodukter som är skadliga för miljön (Jiang et al. 2010) (Zhou et al. 2015) (Lupoi, Gjersing & Davis 2015).

## Vibrationsspektroskopi

I denna studie utnyttjades två typer av vibrationsspektroskopi, Raman och *Fourier Transform Infrared Spectroscopy* (FTIR), i syfte att analysera halten av cellulosa, hemicellulosa, lignin och extraktivämen i veden hos granträ. Båda metoderna utnyttjar molekylers förmåga att vibrera och interagera med ljus. Interaktionen påverkas av ljusets natur och bindningstypen hos molekylen. Från denna data kan ett unikt spektrum av våglängder tillsammans med deras respektive intensitet upprättas utifrån det analyserade provets molekylära sammansättning. Intensiteten hos signalen för en specifik våglängd eller ett intervall av våglängder kan korreleras med den relativa förekomsten av en specifik bindning. Varje bindning vibrerar med en frekvens som styrs av bindningstyp, massa, bindningsstyrka och vibrationstyp. Med bindningstyp menas exempelvis enkel, dubbel eller trippelbindning. Massa syftar på massan hos bindningens ingående atomer i bindningen. Utöver bindningstyp tar bindningsstyrka hänsyn till inflytandet från omgivande grupper. Vibrationstyp syftar till bindningens rörelse i rummet, där de två huvudsakliga kategorierna är sträckning och böjning (Gorzsás 2018). Fördelarna med vibrationsspektroskopi är att metoden är icke-destruktiv, snabb och kräver ingen till relativt lite förbehandling av provet i jämförelse mot traditionella former våtkemiska analyser (Roggo et al. 2007)( Xu et al. 2013).

### FTIR

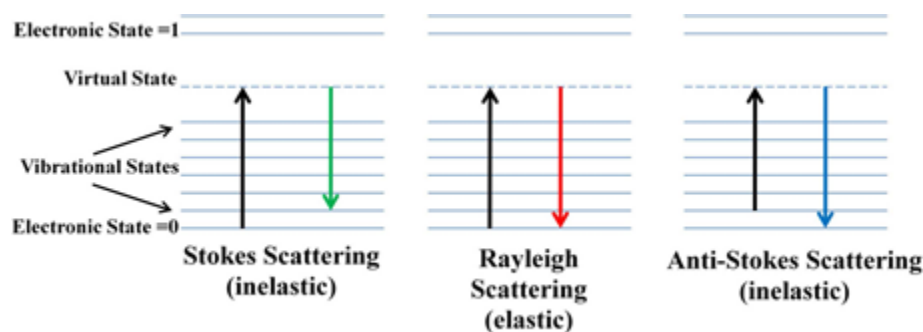
I FTIR utnyttjas molekylers förmåga att absorbera ljus. Absorptionen av vissa våglängder korresponderar till särskilda vibrationsfrekvenser hos dipolmolekyler. En dipolmolekyl är en asymmetrisk fördelning av positiva och negativa laddningar, avståndet mellan laddningarna kallas för dipolmoment (The Penguin Group 1977). Absorptionen korresponderar med en förändring av molekylens dipolmoment (Xu et al. 2013). Således följer att ickepolära molekyler blir transparenta i infrarött ljus. Absorbansen ges utav Beer. Lamberts lag:  $A = a \times l \times c$ , där  $a$  är mediets absorptionsförmåga,  $l$  är sträckan genom mediet, och  $c$  är mediets koncentration (Swinehart 1962). Ett brett omfång av våglängder passerar samtidigt genom provet. Det medför att intensiteten hos varje våglängd mäts under en längre tid men att den totala analys tiden förkortas. För att avlägsna baggrundsabsorption går en referensstråle genom ett transparent medium. Insamlad rådata konverteras till ett spektrum genom en matematisk process kallad *Fourier transformation* (Zhang, Wang & Li 2018). Provet analyseras flera gånger om varav ett genomsnittligt spektrum kan bestämmas. Detta minskar effekten av slumpmässigt brus och förstärker den konstanta signalen från mediet (Zhou et al. 2015).

Det mid-infraröta spektrumet (2500–25,000 nm eller 4000–400  $\text{cm}^{-1}$ ) innehåller våglängder som är centrala för bedömningen av kemiska uppsättningar hos biologiska material (Zhou et al. 2015) (Schwanninger, Rodrigues & Fackler 2011) (Ghaffar & Fan 2013). Inom det mid-infraröda spektrumet finns regioner som förväntas innehålla mer brus (2800-1800  $\text{cm}^{-1}$  och 900-400  $\text{cm}^{-1}$ ) samt absorptionsbandet för vatten (4000-3000  $\text{cm}^{-1}$ ) (Alcantara et al. 2007). Kaliumbromid (KBr) är nästan helt transparent för infraröd strålning och lämpar sig därför som utspädningsmedel och referens vid FTIR. Transparensen hos KBr påverkas negativt av vatten och koldioxid och måste hållas torrt (Zhang, Wang & Li 2018).

### Raman

Inom Ramanspektroskopi utnyttjas molekylers förmåga att sprida inkommande fotoner. Hur stark signal som ges med Raman påverkas av molekylens polariserbarhet, det vill säga dess benägenhet att bli en dipol. När fotonerna träffar mediets molekyler exciteras dessa kortvarigt till ett högre virtuellt vibrationstillstånd, varpå de genast faller till ett stabilt vibrationstillstånd. Från mediet reflekteras en foton som kan ha samma energimängd (kallat

Reylight-spridning eller elastisk spridning), en lägre energimängd (kallat Stokes-våg) eller en högre energimängd (kallat anti-Stokes-våg) än den ingående fotonen (Figur 1). Majoriteten av alla reflekterade fotoner beräknas komma från Reylight-spridning med exakt samma energi/frekvens som den ursprungliga fotonen. Dessa fotoner ger ingen data om mediet och måste filtreras bort från data. En bråkdel ( $\sim 10^{-6}$ ) av de reflekterade fotonerna beräknas vara Stokes-vågor, vilka genereras då molekyler går till ett högre vibrationstillstånd vilket får dem att avge fotoner med lägre energi (lägre frekvens) än de ingående fotonerna. Data samlas vanligen in från differensen i energi mellan den ingående fotonerna och Stokes-vågorna. Anti-Stokes-vågor uppstår när fotonerna träffar molekyler som redan befinner sig i ett exciterat, virtuellt vibrationstillstånd. De avgivna fotonerna från den exciterade molekylen kommer då att ha en högre energi/frekvens än de ingående fotonerna. När anti-Stokes-vågor möter en Stokes-vågor neutraliseras de och försvagar således signalen från mediet. Andelen på förhand exciterade molekyler förväntas dock vara mycket låg vid rumstemperatur (Runcorn et al. 2018) (Lupoi, Gjersing & Davis 2015).



Figur 1. Raman spridning  
 Figure 1. Raman scattering  
 (Lupoi, J.S. 2012 s.18)

Metoden är känsligt för störningar från fluorescerande material, synligt ljus och även kosmisk strålning (Lopez-Reyes & Rull-Pérez 2017). En fördel jämfört med FTIR är att Raman inte störs av vattenmolekyler. Metoden kräver heller inte någon form av förbehandling av provet för att göra analyser, utan fasta material kan skannas direkt. Tidigare har Raman visat sig ha svårigheter med att särskilja hemicellulosa från cellulosa. Pikarna i våglängdspektrumet för hemicellulosa och cellulosa överlappar varandra i Raman vilket gör det svårt att särskilja de olika ämnena åt. Banden för hemicellulosa är svaga och de går att uteslutas vid analyser av cellulosa (Agarwal & Ralph 1997). En orsak till att hemicellulosa och cellulosa är svåra att särskilja är för att hemicellulosa består av huvudgrupper vars kemiska struktur liknar cellulosa, exempelvis xylan (Zeng et al. 2016).

FTIR och Raman utnyttjar sig av två olika fysiska fenomen, absorption och Ramanspridning. Båda fenomenen tillför strålningsenergi som existerar vibrationer inom molekyler (Séne et al. 1994). Däremot ger de upphov till olika spektrum (The Penguin Group 1977). En molekyl ger en stark signal med FTIR om den är kraftigt dipolär, och med Raman om den enkelt kan polariseras. En grov tumregel för detta är att symmetriska bindningar är starkare utslag hos Raman och asymmetriska bindningar ger starkare utslag hos FTIR (Lupoi, Gjersing & Davis 2015).

## Förbehandling av data

Innan insamlad rådata kan användas för att bygga en gemensam modell kan data förbehandlas. I det råa spektrumet finns ofta en andel brus och olikheter i intensiteten som gör

det svårt att direkt använda spektrumet. Dessa skillnader i spektrumet som inte beror på de kemiska sammansättningen försvårar analyser av rådata. Anledningen till dessa skillnader är små avvikelser i proverna och miljö. Det kan vara i form av att ytan på proverna inte är helt jämna vilket ger brus eller att mängden i varje prov varierar vilket ger olika intensitet. För att då få data som kan analyseras krävs ofta att de olika spektrumen förbehandlas med en gemensam baslinje, utjämning och en normalisering. Detta ger analyserbara data där skillnader skapade av brus och olika intensitet har tagits bort (Felten et al. 2015).

### ***Baslinje***

En gemensam baslinje innebär att formen hos de olika spektrumen placeras på en gemensam axel. För biologiska material bestäms baslinjen lämpligen med hjälp av *Asymmetrical Least Squares* (AsLS). Baslinjen justeras utifrån två parametrar;  $\lambda$  och P. Första parametern  $\lambda$  avgör till vilken grad som baslinjen följer spektrumens form, där ett högt  $\lambda$  innebär en linjär baslinje och ett lågt  $\lambda$  innebär att linjen perfekt följer spektrumens form. Andra parametern P, avgör det tillskrivna vikten hos positiva residualer i spektrumet. För spektroskopi av biologiska material blir P ointressant eftersom spektrumet inte förväntas innehålla negativa värden (Felten et al. 2015).

### ***Utjämning***

Utjämning av data utnyttjas för att minska effekten av slumpmässigt brus som uppstår vid analysen. Vid för stor utjämning riskerar dock spektrumet att förlora skillnader som faktiskt är signifikanta. Den mest använda metoden, Savitzky-Golay filtrering (S-G), består utav två parametrar; polynom ordning och ramstorlek. Desto större differens mellan polynom ordning och ramstorlek desto större blir utjämningen av data (Felten et al. 2015).

### ***Normalisering***

Normalisering av spektrumen innebär att intensiteten skalas om utifrån särskilda villkor. Detta kan göra dem enklare att visualisera. Det är inte nödvändigt för att kunna analysera spektrumen men det kan ge en tydligare modell. Det kan också vara viktigt för att vid exportering av spektrumen kunna bearbetas med andra verktyg. För normalisering finns flera alternativ, bland annat TotArea, Regionminmax och Regionmax. TotArea innebär att varje intensitet divideras med summan av alla ingående intensiteter. Regionminmax innebär att den högsta intensiteten inom en utvald region av spektrumen sätts till värde 100 och den lägsta intensiteten i regionen till värde 0. Övriga intensiteter skalas i relation till dessa värden. RegionMax innebär att inom en utvald region av spektrumen sätts den högsta intensiteten till värde 100 och den lägsta intensiteten i hela spektrumen till värde 0. Övriga intensiteter skalas i relation till dessa värden (Felten et al. 2015).

## **Modellering**

För att skapa modeller för spektraldata krävs statistiska verktyg som kan hantera många prediktioner för att förutsäga en respons. I kontexten av denna studie är prediktionerna intensiteten hos alla de våglängder som modellen använder och responsen är halten av en kemisk komponent. Detta förutsätter alltså att responsen redan är känd för att kunna bygga en modell. *Partial least squares* (PLS) procedur är en form av multivariansanalys och ett lämpligt verktyg för att modellera spektrum (Chen Ferrari & Bramanti 2010). Den bygger på att skapa nya så kallade principiella faktorer utifrån kombinationer av insamlade variabler, i syfte att de maximera förklaringsgraden hos responsen. Fler faktorer gör att modellen bättre kan förutsäga responsen för tillgänglig data. För många faktorer kan dock innebära att modellen blir sämre på att förutsäga responsen hos framtida data, kallat överanpassning



(*overfitting*) (Mannan & McNeil 2013). Fler faktorer medför också att modellen blir mer komplex, vilket bör undvikas.

Modellens styrka kan bedömas utifrån dess förmåga att förutsäga responsen hos observationer som inte ingick vid dess konstruktion. Data kan därför delas upp i två grupper. Första gruppen används som träningsgrupp för att bygga modellen. Den andra gruppen används som testgrupp för att utvärdera modellen. Metoden kallas korsvalidering, (CV - *cross validation*) och är starkt beroende av provstorlek för att skapa stora tränings- och testgrupp som möjligt. Korsvalidering utnyttjas för att välja ett optimalt antal principiella faktorer. Följande process upprepas: uppdelning av tränings- och testgrupp, konstruktion av principiella faktorer baserad på träningsgruppen för att förutsäga så mycket som möjligt av responsen i form av den halten en kemisk komponent i testgrupp. För att bestämma hur många faktorer som är optimalt för modellen adderas kvadraten av alla differenser mellan det predikterade och det observerade värdet för att beräkna ett PRESS-värde (*predicted residual error sum of squares*). Modellens PRESS-värde beräknas för varje ny faktor och är ett mått på modellens prediktiva styrka (Chen, Ferrari & Bramanti 2010). Det maximala antalet faktorer för modellen baseras på det antal som minimerar PRESS-värdet. För att minimera antalet faktorer ytterligare extraheras endast de som ger en signifikant förbättring av PRESS. Notera att de faktorer som i slutändan extraheras inte nödvändigtvis ger den högsta förklaringsgraden (SAS Institute Inc 2008).

## Syfte

Syftet med denna rapport är att utvärdera och jämföra FTIR och Raman, som potentiella metoder för att bestämma den kemiska sammansättningen hos granved. Detta kombinerat med olika typer förbehandlings, vilket för FTIR innebär olika former av normaliseringar och för Raman olika provmaterial. För Raman förväntas träpulverprover prestera bättre än borkärnor som provmaterial. Metoderna utnyttjar sig utav olika fysiska fenomen för att samla in data och förväntas därför kunna komplettera varandra. FTIR förväntas prestera bättre vid modellering av hemicellulosa. Raman förväntas prestera bättre vid modellering av lignin.

## **MATERIAL & METOD**

### **Förbehandling**

Fem D-formade borrhärdar av gran, slipade på en sida, användes i studien. Varje borrhärd delades upp i tre lika stora regioner att analyseras; en närmast kärnan, en i mitten och en närmast kambiet. Bitar av borrhärdan i varje region skars bort. Proverna pulveriserades i en askberedare i 2 x 40 sekunder med två minuter mellan sessionerna. 1-10 mg av träpulvret blandas med 399-390 mg (KBr) och maldes med en mortel till en mjölig struktur. Totalt användes 15 prover från 5 borrhärdar.

### **Kemisk analys**

Från MoRe Research i Örnsköldsvik hämtades information om de 5 borrhärdarnas halter av arabinos, galaktos, glukos, xylos, mannos, cellulosa, hemicellulosa, lignin och extraktivämnen samt andelen av borrhärdarnas totala massa som kunde bestämmas. Halten hemicellulosa baseras på summan av arabinos, galaktos, xylos och mannos. På grund av den låga provstorlekar gjordes justeringar för analysmetoderna. Sockermolekyler analyserades enligt metodkod SCAN-CM 71:09, med modifikationen att endast ett prov på 100 mg användes per borrhärd (istället för två prover med 300 mg var). Ligninhalten bestämdes utifrån två standardanalysen kallade Klason lignin och syralösligt lignin. För Klason lignin användes en metod som liknar metodkod Tappi T222 med modifikationen att potentiellt störande extraktivkemikalier inte kunde extraheras innan ligninhalten uppmättes. Istället beräknades ligninhalten på följande vis:

$$\text{Lignin Klason} = \text{Lignin (uppmätt)} - \text{extraktivämnen}.$$

Syralösligt lignin analyserades enligt standard utifrån metodkod Tappi T-UM 250. Totala ligninhalten beräknades som summan av Klason lignin och syralösligt lignin.

Extraktivämnen analyserades utifrån en "hemmagjord" metod med hjälp av en Soxtechextraktor, varav 100 mg av proverna användes (istället för 5 g).

Tabell 1. Kemisk sammansättning hos de fem borrhärnorna 1-5, uttryckt i % av massan. Borrhärnorna 1-5 är uppdelade i tre regioner; 1= mörk, 2= mitt och 3=bark. Summan baseras på den sammanlagda massan hos extraktivämnen, lignin, cellulosa och hemicellulosa.

*Table 1. Chemical composition in the five wood core 1-5 explain in 5 of the mass. The wood core 1-5 is divided in three regions; 1=pith 2=middle and 3=bark. The sum is based on the total mass from extractives, lignin, cellulose and hemicellulose*

Prov Nr	1_1	1_2	1_3	2_1	2_2	2_3	3_1	3_2
Arabinos	1,24	0,93	0,94	0,13	0,97	0,79	1,5	0,92
Galaktos	2,09	1,81	2,11	2,96	4,08	3,93	4,16	1,6
Glukos	42,2	44,7	43	40,1	42,4	42,4	38,6	45,1
Xylos	6,08	5,19	4,34	5,36	5,04	4,59	5,08	4,86
Mannos	9,49	9,68	9,89	10,4	9,25	9,46	9,65	9,94
Extraktivämnen	1,2	0,9	1,9	1,4	1,3	1,5	1,9	1,2
Lignin	28,5	28,1	27,4	27,7	30,7	31,4	30,8	28
Cellulosa	39,04	41,5	39,72	36,61	39,33	39,27	35,39	41,76
Hemicellulosa	22,06	20,81	20,56	23,51	22,47	21,9	23,65	20,66
Sum : (cell + hem + lig)	90,8	91,31	89,58	89,22	93,74	94,07	91,69	91,62
Prov Nr	3_3	4_1	4_2	4_3	5_1	5_2	5_3	
Arabinos	0,93	1,32	1,2	1,08	1,29	0,85	0,91	
Galaktos	1,48	2,23	1,97	1,24	2,52	1,61	1,88	
Glukos	47,7	41,5	44,1	46,3	41,5	45,4	45,5	
Xylos	4,39	5,21	4,85	4,59	6,09	5,14	4,84	
Mannos	9,93	9,66	10,9	9,36	9,77	10,6	9,77	
Extraktivämnen	1,2	1,2	1,1	1,4	1,1	1,5	2,2	
Lignin	26,6	30,2	30,1	28,5	30,3	26,7	27	
Cellulosa	44,39	38,29	40,52	43,17	38,23	41,91	42,27	
Hemicellulosa	20,04	21,63	22,5	19,4	22,94	21,69	20,69	
Sum: (cell + hem + lig)	92,23	91,32	94,22	92,47	92,57	91,8	92,1	

Tabell 2. Statistisk diagnostik för den kemiska sammansättningen för de 15 träproverna.

*Table 2. Statistical diagnostics for the chemical composition for the 15 wood Samples*

(%)	Medel	Min	Max	SD
Arabinos	1,08	0,79	1,5	0,21
Galaktos	2,38	1,24	4,16	0,93
Glukos	43,37	38,6	47,7	2,38
Xylos	5,04	4,34	6,09	0,5
Mannos	9,85	9,25	10,9	0,45
Extraktivämnen	1,4	0,9	2,2	0,34
Lignin	28,8	26,6	31,4	1,58
Cellulosa	40,09	35,39	44,39	2,37
Hemicellulosa	21,63	19,4	23,65	1,22
Sum: (lig+hem+cell)	91,92	89,22	94,22	1,39

## FTIR-skanning

FTIR-spektrumet togs fram med en IFS 66v/S spektrofotometer. För bakgrund maldes 400 mg ren kaliumbromid(KBr) med mortel. Proverna med malt träpulver och KBr placerades i stålpatroner och analyserades vid 0,007 atmosfäriskt tryck. Varje prov skannades 128 gånger med en upplösning av 4  $\mu\text{m}$ . Rådata konverterades till spektrum i programvaran OPUS v.5.5 som en XPM-fil. Erhållna spektrum exporterades till OPUS v 7.8 för att kunna exporteras vidare som en MATLAB data-fil.



Bild 1. IFS 66v/S spektrofotometer (Sally). Foto: Edward Carlsson.

## RAMAN-förbehandling

Samma 5 borrhärnor användes för Raman. En del av det träpulver som gavs vid förberedelserna för FTIR proverna sparades till RAMAN och blandades inte med KBr. De borrhärnor som träpulverproverna togs från var alla försedda med en plan och slipad yta som lämnades intakt.

## RAMAN-skanning

Två typer av granprover användes för Ramanspektroskopi: 5 slipade borrhärnor (samma som vid FTIR) där dess tre regioner analyserades separat, och dess 15 korresponderande träpulverprover placerade i små glastuber. Varje prov hölls fast framför lasern. Ramanspektrumet togs fram med hjälp av en handhållen BRAVO Ramanspektrometer. Skanningarna utfördes i ett mörklagt rum för att minimera brus från andra ljuskällor än lasern. Spektrumet för varje prov skannades mellan 300-3200  $\text{cm}^{-1}$ . Standardinställningar för automatisk analys med användes. Rådata konverterades till spektrum i programvaran OPUS v7.8 som en XPM-fil. Därefter exporterade samtliga spektrum som en gemensam MATLAB data-fil.



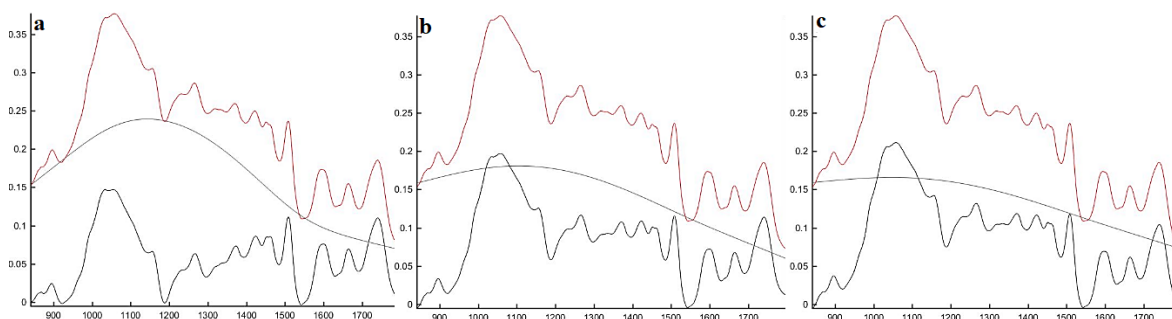
Bild 2. BRAVI Ramanspektrometer. Foto: Edward Carlsson.

## Förbehandling av data

MATLAB data-filen för öppnades i Matlab 9.0.0, där spektrumen från träpulverproverna och borrhärnproverna separeras till två separata filer. Scriptet MCR-ALS GUI v4c användes som mall för förbehandling av dessa spektrum.

## FTIR

Våglängderna avgränsades till  $840,8\text{--}1\,799,3\text{ cm}^{-1}$ . För val av optimal baslinjen gjordes tre tester med  $P = 0,001$  och  $\lambda$ -värde: 50 000, 500 000 och 1 000 000 (Figur 2) med målet att minimera PRESS i SAS. Alla tre baslinjer uppvisade samma PRESS-värde men  $\lambda$ -värdet sattes till 1 000 000 för att minimera risken för överanpassning. För utjämning användes standardinställningar (polynom ordning 1 och ramstorlek 3). För normalisering användes Total area. På rekommendation av Andra Gorzsas, forskningsingenjör vid Umeå Universitet, testades även ytterligare sex normaliseringar: Regionmax ( $925\text{--}1145\text{ cm}^{-1}$ ), Regionmax ( $1487\text{--}1553\text{ cm}^{-1}$ ), Regionmax ( $1553\text{--}1633\text{ cm}^{-1}$ ), Regionminmax ( $925\text{--}1145\text{ cm}^{-1}$ ), Regionminmax ( $1487\text{--}1553\text{ cm}^{-1}$ ) och Regionminmax ( $1553\text{--}1633\text{ cm}^{-1}$ ). Ett dataset skapades utan normalisering. Våglängdsregionerna, illustrerade i Figur 3, valdes utifrån rekommendationer från Tomas Funda (postdoc vid SLU - institutionen för skoglig genetik och växtfysiologi).



Figur 2. Baslinjer för ett FTIR spektrum. Översta linje visar formen hos det ursprungliga spektrumet, mellersta linjen visar baslinjen, och nedersta linjen visar resulterande spektrum efter baslinjekorrigering. **a)**  $\lambda = 50\,000$ . **b)**  $\lambda = 500\,000$ . **c)**  $\lambda = 1\,000\,000$ .

*Figure 2 Baselines for a FTIR spectrum. Top line shows the shape of original spectrum, middle line shows the*

baseline, and the bottom line shows resulting spectrum after baseline correction. **a)**  $\lambda = 50\,000$ . **b)**  $\lambda = 500\,000$ . **c)**  $\lambda = 1\,000\,000$ .

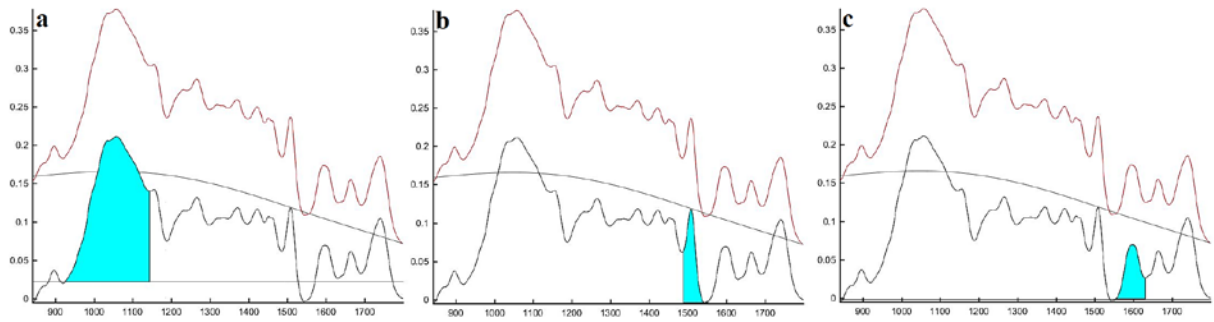


Figure 3 Olika regioner för normalisering. **a)**  $925-1145\text{ cm}^{-1}$ . **b)**  $1487-1553\text{ cm}^{-1}$ . **c)**  $1553-1633\text{ cm}^{-1}$ .

Figure 3. Different regions for normalization. **a)**  $925-1145\text{ cm}^{-1}$ . **b)**  $1487-1553\text{ cm}^{-1}$ . **c)**  $1553-1633\text{ cm}^{-1}$ .

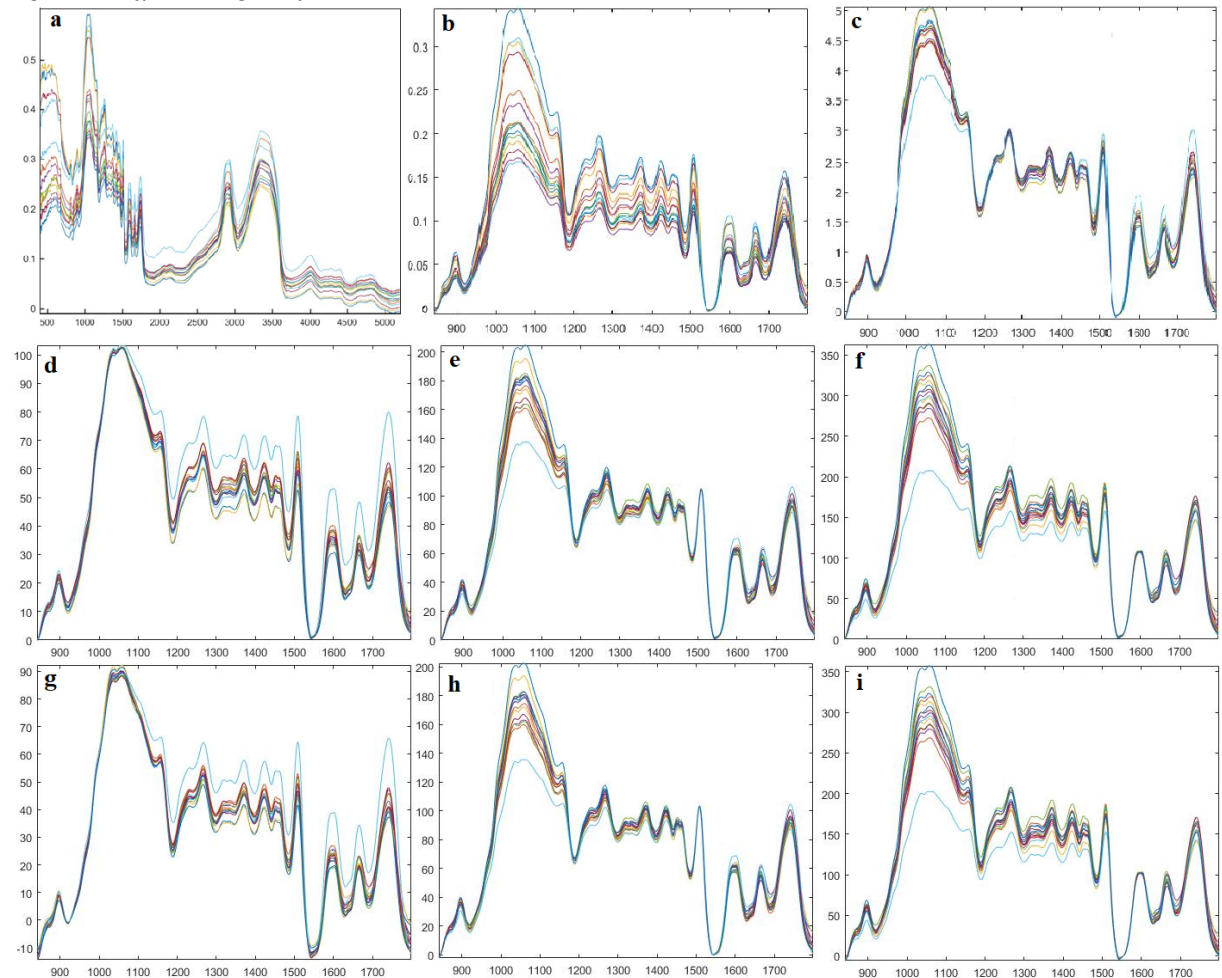


Figure 4. Jämförelse mellan förbehandlade och icke förbehandlade FTIR-spektrum. **a)** Visar icke förbehandlade FTIR-spektrum. **b)** Visar det spektrum med våglängderna avgränsade till  $840,8-1\,799,3\text{ cm}^{-1}$  och baslinje på  $\lambda$ -värde  $1\,000\,000$  som användes vid normaliseringen. **c-i)** visar de normaliserade spektrum **c)** visar Total area som normalisering. **d-f)** visar Regionmax normalisering med regionerna  $925-1125\text{ cm}^{-1}$ ,  $1487-1553\text{ cm}^{-1}$  och  $1553-1633\text{ cm}^{-1}$ . **g-i)** visar Regionminmax normalisering med regionerna  $925-1125\text{ cm}^{-1}$ ,  $1487-1553\text{ cm}^{-1}$  och  $1553-1633\text{ cm}^{-1}$ .

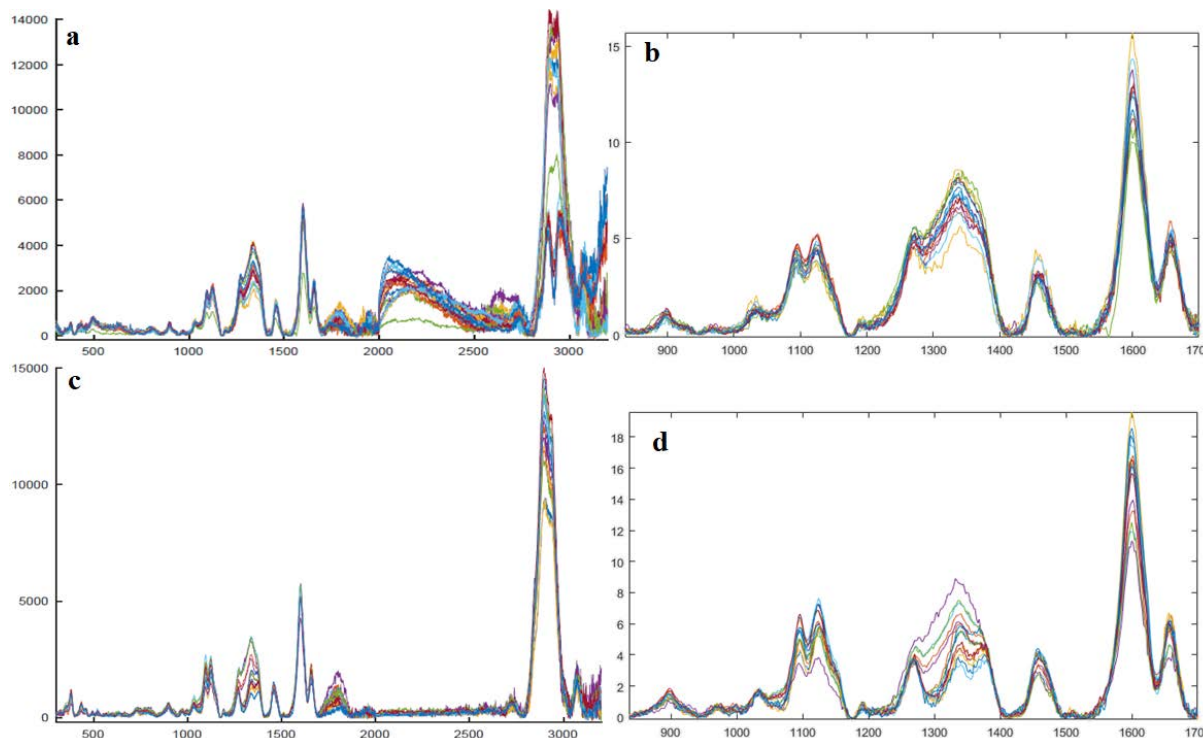
Figure 4. Comparison between the untreated and pre-treated FTIR spectra. **a)** Displays the untreated FTIR spectra. **b)** Displays the spectra with wavelengths delimited to  $840.8-1\,799.3\text{ cm}^{-1}$  and baseline at  $\lambda$  value  $1\,000\,000$  used in normalization. **c-i)** shows the normalized spectra with **c)** showing total area as normalization. **d-f)**



shows Regionmax normalization with the regions  $925\text{--}1125\text{ cm}^{-1}$ ,  $1487\text{--}1553\text{ cm}^{-1}$  and  $1553\text{--}1633\text{ cm}^{-1}$ . **g-i)** shows Regionminmax normalization with the regions  $925\text{--}1125\text{ cm}^{-1}$ ,  $1487\text{--}1553\text{ cm}^{-1}$  and  $1553\text{--}1633\text{ cm}^{-1}$ .

## Raman

Våglängderna avgränsades till  $836\text{--}1700\text{ cm}^{-1}$ . Samma baslinje och utjämning som för FTIR användes ( $P=0,001$  och  $\lambda = 1\,000\,000$ ). Två dataset skapades, ett utifrån borrhärdarna och ett utifrån träpulvret. För båda dataseten användes Total area som normalisering.



Figur 5. Jämförelse mellan de icke förbehandlade ( $300\text{--}3200\text{ cm}^{-1}$ ) och förbehandlade ( $836\text{--}1700\text{ cm}^{-1}$ ) Raman spektrum. **a)** och **b)** visar de icke förbehandlade och förbehandlade spektrum för Raman med träpulver. **c)** och **d)** visar de icke förbehandlade och förbehandlade spektrum för Ramanproverna med borrhärdar.

Figure 5. Comparison of the untreated ( $300\text{--}3200\text{ cm}^{-1}$ ) and pretreated ( $836\text{--}1700\text{ cm}^{-1}$ ) Raman spectra. **a)** and **b)** shows the difference between the untreated and the pre-treated Raman spectra of the samples with wood powder. **c)** and **d)** Shows the difference between the untreated and the pre-treated Raman spectra of the samples with wood core.

## Excel

De tre förbehandlade dataseten, 1 FTIR (malt träpulver med 8 normaliseringar) och 2 Raman (träpulver och borrhärdar) exporterades till ett Excel ark tillsammans med information från den kemiska analysen för de 15 proverna. Extra spektrum som uppstod vid test skanningar och felmätningar togs bort. Resterande spektrum kopplades till sina respektive kemiska referensvärden. Excelarken sparades som 10 txt.-filer.

## SAS

Programvaran SAS v 9.4 användes för att skapa och utvärdera modeller av provernas kemiska sammansättning. Txt.-filen för respektive dataset importerades till SAS där analysen utav alla dataset utfördes utifrån varianter av följande script:

```
proc pls data=« DATA » cv=split cvtest(stat=PRESS pval=0.1) nfac=25 plot=(diagnostics dmod);  
model «VARIABLE» = « _XXX_X » - « _XXX_X »;  
run;
```

**proc pls** = specificerar att PLS-proceduren ska användas.

**DATA** = specificerar namnet på txt.-filen för valt dataset.

**cv=split** = specificerar att korsvalidering skall utföras och vart *n:e* prov ska ingå i testgruppen. Förinställningen för programmet är att var 7:e prov väljs ut.

**cvtest** = specificerar att van der Voets slumpbaserade modelljämförelsetest används för att jämföra ifall skillnaden mellan olika modeller är statistiskt signifikant.

**(stat=PRESS)** = specificerar att vid jämförelsen för **cvtest** ska baseras på modellernas PRESS-värde.

**pval** = specificerar för **cvtest** hur stor skillnaden är mellan två faktorer för att den ska anses vara signifikant. Förinställning för programmet är att  $pval = 0,1$ .

**nfac** = Högsta antal faktorer som ska skapas.

**plot=(diagnostics dmod)** = specificerar att i resultatet visas diagram över potentiella avvikande värden.

**model** = specificerar den modell som ska testas.

**VARIABLE** = den utvalda responsfaktor som modellen försöker förutsäga.

**XXX\_X – \_XXX\_X** = det utvalda våglängdsintervallet från vilka värden tas för att förutsäga responsvariabeln.

Alla tre dataset, (8 FTIR - malt träpulver, 1 Raman - borrhärnareprover och 1 Raman - träpulverprover) användes för att skapa modeller för halten av; arabinos, galaktos, glukos, xylos, mannos, cellulosa, hemicellulosa, lignin och extraktivämnen. Vid korsvalideringen användes 13/15 prover till träningsgrupp och 15/2 prover till testgrupp. För varje modell noterades: lägsta PRESS-värde, antal faktorer som gav upphov till lägsta PRESS-värde och antal faktorer ur föregående kategori som sänkte PRESS-värde med minst 0,1. Först testades alla dataset med  $pval = 0,1$  (standard). De modeller vars lägsta PRESS-värde inkluderade fler än noll faktorer valdes ut för en andra omgång tester utan minimikrav på  $pval$ . Denna gång optimerades modellerna ytterligare genom att ta bort tydligt avvikande värden från samtliga modeller utifrån i syfte att öka dess  $R^2$  utan att överanpassa den eller öka dess ursprungliga PRESS-värde. Vad som räknades som avvikande värden bedömdes visuellt utifrån diagram i SAS.



# RESULTAT

## FTIR

Från testomgången med FTIR-spektrumen med krav på att  $pval=0,1$  kunde SAS inte inkludera principiella faktorer för alla modeller (Tabell 3). Endast modellerna för arabinos och mannos fick ett lägsta PRESS-värde där faktorer kunde inkluderas. För lignin kunde endast en faktor inkluderas vid lägsta PRESS-värde då ingen normalisering utfördes. Ingen av faktorerna hos arabinos, mannos eller lignin ansågs dock medföra en signifikant förbättring av modellen och extraherades därför inte. Eftersom ingen faktor kunde extraheras kunde SAS inte redovisa  $R^2$  för någon av modellerna.

Tabell 3. PRESS och antal principiella faktorer för FTIR-modeller, med  $pval = 0,1$

Table 3. PRESS and for Number of principal factors FTIR models, with  $pval = 0,1$

Kemiska komponenter	Arabinos	Galaktos	Glukos	Xylos	Mannos	Extraktiv	Lignin	Cellulosa	Hemicellulosa
PRESS									
Regionmax 925-1145	1,0842	1,124	1,1319	1,0596	1,0737	1,0674	1,1413	1,138	1,1611
Region max 1487-1553	1,0137	1,124	1,1319	1,0596	1,1199	1,0674	1,1413	1,138	1,1611
Region max 1553-1633	1,0914	1,124	1,1319	1,0596	1,0066	1,0674	1,1413	1,138	1,1611
Regionminmax 925-1145	1,0986	1,124	1,1319	1,0596	1,1969	1,0674	1,1413	1,138	1,1611
Regionminimax 1487-1553	0,9704	1,124	1,1319	1,0596	1,1682	1,0674	1,1413	1,138	1,1611
Regionminmax 1553-1633	1,089	1,124	1,1319	1,0596	1,0981	1,0674	1,1413	1,138	1,1611
Total area	0,9594	1,124	1,1319	1,0596	1,1969	1,0674	1,1413	1,138	1,1611
Ingen normalisering	1,061	1,124	1,1319	1,0596	1,194	1,0674	1,0673	1,138	1,1611
Minsta antal faktorer									
Regionmax 925-1145	2	0	0	0	5	0	0	0	0
Region max 1487-1553	2	0	0	0	5	0	0	0	0
Region max 1553-1633	1	0	0	0	4	0	0	0	0
Regionminmax 925-1145	0	0	0	0	0	0	0	0	0
Regionminimax 1487-1553	2	0	0	0	5	0	0	0	0
Regionminmax 1553-1633	1	0	0	0	4	0	0	0	0
Total area	2	0	0	0	0	0	0	0	0
Ingen normalisering	3	0	0	0	5	0	1	0	0

När FTIR-spektrumen testades utan krav på  $pval$  kunde förklaringsgraden ( $R^2$ ) redovisas. Från Tabell 4 framgår det att fler extraherade faktorer generellt gav ett högre  $R^2$ . När avvikande värden togs bort höjdes  $R^2$ . Detta gav dock inget ingen signifikant förbättring för arabinosmodellerna. Även när en faktor extraherades, som för datasetet utan normaliseringen eller två faktorer som för Regionmax 925-1125, skapades ingen modell med tillfredsställande  $R^2$ . Modeller för mannos uppvisade generellt en hög förklaringsgrad oavsett om några avvikande värden exkluderades. Med Regionmax 1487-1553 och Regionmax 1553-1633 kunde en faktor exkluderas med resultatet att modellens  $R^2$  vart högre. För Regionminmax 1553-1633 och utan någon normalisering kunde två faktorer exkluderas med resultatet att  $R^2$  vart högre. Det framkom att samma avvikande värden ofta återkom från de tre proverna; 1\_1, 1\_3 och 3\_2.

Tabell 4. PRESS, antal principiella faktorer, extraherade faktorer och förklaringsgrad av responsten ( $R^2$ ) för FTIR-modeller utan krav på pval. De blåmarkerade rutorna indikerar att minst en faktor extraherades. \*) ett avvikande värde togs bort. \*\*) två avvikande värden togs bort

Table 4. PRESS, Number of principal factors, Number of extracted factors and Coefficient of determination ( $R^2$ ) for FTIR models without demands on pval. The blue marked boxes indicate that one factor was extracted. \*) one outlier was removed. \*\*) two outliers were removed

Kemiska komponenter	Arabinos	Mannos	Lignin	Kemiska komponenter	Arabinos	Mannos	Lignin
PRESS				extraherade faktorer			
Regionmax 925-1145	1,0759**	1,0737	-	Regionmax 925-1145	2**	5	-
Region max 1487-1553	1,0137	1,0374*	-	Region max 1487-1553	2	5*	-
Region max 1553-1633	1,0914	0,949*	-	Region max 1553-1633	1	4*	-
Regionminimax 1487-1553	0,9704	1,1682	-	Regionminimax 1487-1553	2	5	-
Regionminimax 1553-1633	1,089	0,7624**	-	Regionminimax 1553-1633	1	4**	-
Total area	0,9594	-	-	Total area	2	-	-
Ingen normalisering	0,8644*	0,9036**	0,9155*	Ingen normalisering	3*	5**	2*
Minsta antal faktorer				$R^2$			
Regionmax 925-1145	2**	5	-	Regionmax 925-1145	64,6483**	93,722	-
Region max 1487-1553	2	5*	-	Region max 1487-1553	41,1459	93,047*	-
Region max 1553-1633	1	4*	-	Region max 1553-1633	16,0092	92,4133*	-
Regionminimax 1487-1553	2	5	-	Regionminimax 1487-1553	42,442	91,7	-
Regionminimax 1553-1633	1	4**	-	Regionminimax 1553-1633	16,45	95,7597**	-
Total area	2	-	-	Total area	44,8606	-	-
Ingen normalisering	3*	5**	2*	Ingen normalisering	63,0537*	86,1439**	61,5006*

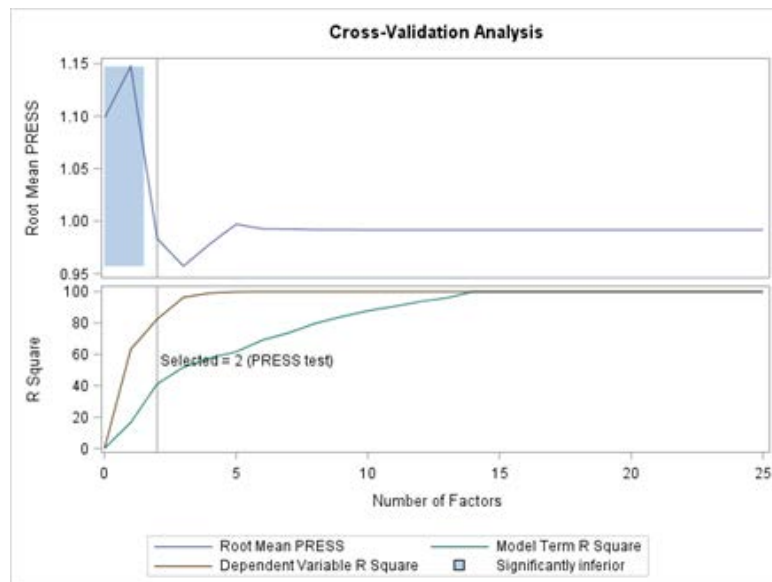
## Raman

Från testomgången med Raman-spektrumen med krav på att pval = 0,1 kunde SAS inte inkludera principiella faktorer för alla modeller (Tabell 3). Ingen av modellerna för galaktos, mannos eller lignin fick ett lägsta PRESS-värde där principiella faktorer var inkluderade. Datasetet från borrhärdarna kunde faktorer inkluderas vid lägsta PRESS-värde för glukos, xylos extraktivämnen, cellulosa och hemicellulosa. Från datasetet baserat på träpulvret kunde faktorer inkluderas vid lägsta PRESS-värde för arabinos, glukos, cellulosa och hemicellulosa. Ingen av faktorerna ansågs vara signifikanta, med undantag för modellen för arabinos som extraherade två signifikanta faktorer med ett  $R^2$  på 82,51 %. En tredje faktor kunde inkluderas för arabinos men denna var inte signifikant. PRESS-värdet för arabinos var lägre med noll faktorer än med endast 1 faktor (Figur 6).

Tabell 5. PRESS och antal principiella faktorer för Ramanmodeller med träpulverproverna och borrhärdproverna, med pval = 0,1

Table 5. PRESS and Number of principal factors( $R^2$ ) for Raman-models with wood powder samples and wood core samples, with pval = 0,1

Kemiska komponenter	Arabinos	Galaktos	Glukos	Xylos	Mannos	Extraktiv	Lignin	Cellulosa	Hemicellulosa
PRESS									
Borrhärna	1,0986	1,124	1,1252	1,0438	1,1969	1,0247	1,1413	1,1333	1,1467
Pulver	0,9571	1,124	1,0815	1,0596	1,1969	1,0674	1,1413	1,0697	1,1433
Minsta antal faktorer									
Borrhärna	0	0	1	1	0	4	0	2	4
Pulver	3	0	2	0	0	0	0	2	4



Figur 6. Diagram från SAS för korsvalidering av arabinos från Ramanspektrumen från träpulverproverna med  $pval = 0,1$ .

Figure 6. Diagram from SAS for cross validation of arabinose from Rama spectra from the wood powder samples with  $pval = 0,1$ .

När Raman-spektrumen testades utan krav på  $pval$  kunde förklaringsgraden  $R^2$  redovisas. Från tabell 6 framgår det att fler extraherade faktorer generellt gav ett högre  $R^2$ . Från datasetet med träpulverproverna gick det inte att exkludera några avvikande värden utan att modellerna blev fullkomligt överanpassad, med undantag för cellulosa, där en faktor kunde exkluderas. Modeller för arabinos, hemicellulosa och extraktivämnen uppnådde alla  $R^2$  på över 95 % utan att något avvikande värde togs bort. Glukos och xylos gav för borkärnproverna inga modeller med ett tillräckligt högt  $R^2$  för att anses vara tillfredsställande. Det gick inte heller att förbättra modellerna genom att exkludera några avvikande faktorer. Skillnaden i modellernas styrka var särskilt påtaglig för glukos, där borkärnproverna gav ett  $R^2$  på 27,3 % och träpulverproverna 92,9 %. Det framkom att samma avvikande värden ofta återkom från de tre proverna; 1\_2, 3\_1 och 3\_3. Dessa tre prover återkom ofta samlade vilket gjorde de inte gick att exkludera.

Tabell 6. PRESS, antal principiella faktorer, extraherade faktorer och förklaringsgrad av responsen ( $R^2$ ) för Raman-modeller med träpulverproverna och borrhänsproverna, utan krav på pval. De blåmarkerade rutorna indikerar att minst en extraherades. \*) ett avvikande värde togs bort

Table 6. PRESS, Number of principal factors, Number of extracted factors and Coefficient of determination ( $R^2$ ) for Raman-models with wood powder samples and wood core samples, without demands on pval. The blue marked boxes indicate that one factors was extracted. \*) one outlier was removed

Kemiska komponenter	Arabinos	Glukos	Xylos	Extraktiv	Cellulosa	Hemicellulosa
PRESS						
Borrhäns	-	1,1252	1,0438	1,0247	1,1333	1,1467
Pulver	0,9571	1,0815	-	-	1,0833	1,1433
Minsta antal faktorer						
Borrhäns	-	1	1	4	2	4
Pulver	3	2	-	-	2	4
extraherade faktorer						
Borrhäns	-	1	1	4	2	4
Pulver	3	2	-	-	2	4
$R^2$						
Borrhäns	-	27,322	35,6671	98,6423	72,2902	96,2974
Pulver	96,3606	92,9166	-	-	88,9913*	99,4855

## DISKUSSION

SAS lyckades endast skapa en modell utifrån kraven  $p_{val} = 0,1$  när spektrumen från träpulverproverna i Raman användes för att förutsäga arabinos. Två faktorer kunde då extraheras, varav den första förklarade 63,3 % av variationen och den andra 19,2 %. Utan krav på  $p_{val}$  kunde även en tredje faktor extraheras vilket då ökade  $R^2$  till 96 %. Detta ansågs dock vara på gränsen till överanpassning (Tabell 5). Varför just arabinos lyckades i första omgången var inte självklart. En möjlig förklaring skulle vara att arabinos hade den lägsta variationen av alla kemiska komponenter (Tabell 2). Detta kan således ha medfört att värdet blev enklare att förutsäga. Detta förklarar dock ej varför FTIR eller Ramanmodellerna från borrhärneproverna inte presterade lika bra. Det är värt att notera att med enbart den första faktorn blev PRESS högre för arabinos modellen jämfört med om inga faktorer extraherades (Figur 6).

Eftersom SAS endast inkluderade faktorer för arabinos, mannos och lignin vid lägsta PRESS-värde från FTIR kunde endast dessa jämföras efter att krav på  $p_{val}$  tagit bort. FTIR-modellerna för arabinos utan krav på  $p_{val}$  presterade väldigt olika beroende på typ av normalisering, då skillnaden hos  $R^2$  mellan den bästa och sämsta modellen var 48 % (Tabell 4). De två sämsta modellerna tycktes avvika från de övriga, med  $R^2$  på 16,0 % och 16,5 %. Intressant var att de båda använde sig av regionen  $1553 - 1633 \text{ cm}^{-1}$  för normalisering. Övriga arabinosmodeller presterade avsevärt bättre, med ett  $R^2$  på 41,1 - 64,6 %. Detta indikerar att normalisering kunde ha en betydande effekt på modellens styrka och mer specifikt att regionen  $1553 - 1633 \text{ cm}^{-1}$  är mindre lämplig för att förutsäga arabinos. För mannos presterade alla FTIR-modeller väldigt bra, med regelbundet höga  $R^2$ . Typen av normalisering spelade alltså en mindre roll för att förutsäga mannos. Den enda gång som FTIR inkluderade någon faktor för lignin vid lägsta PRESS-värde var då ingen normalisering användes. Eftersom de resterade sex modellerna misslyckades och den skapade modellen hade ett lågt  $R^2$  anses ändå FTIR ha presterat dåligt för lignin.

En korrekt utförd förbehandling av proverna inverkar på spektrumens kvalitet (andel brus). Detta eftersom de fysiska fenomen som utnyttjas av FTIR och Raman påverkas provets mekaniska egenskaper. Både absorption och Ramanspridning påverkas av bland annat mediets porositet, ytegenskaper och densitet (Schwanninger Rodrigues, & Fackler 2011). Texturen hos träpulvret kontrollerades därför flera gånger under övervakning av Tomas Funda. De resulterande FTIR-spektrumen bedömdes visuellt innehålla en låg andel brus inom våglängdsintervallet av intresse (Figur 4). Även dessa kontrollerades av Tomas Funda. Således bedömdes förbehandlingen (genom manuell skärning, malning och användning av askberedare) ha varit tillfredsställande.

Spektrum från Raman bedömdes visuellt innehålla betydligt mer brus än FTIR (Figur 4 & 5). Detta var väntat, då Ramanspridning antas vara ett mer oberäkneligt fysiskt fenomen att mäta än absorption. Den miljondel av alla fotoner som bildar Stokes-vågor kan inte förväntas flyga iväg i samma riktning varje gång, utan detta beror till stor del på jämnheten i mediets yta. Signalstyrkan från mediet hos Raman förväntas bli exponentiellt svagare med ökat avstånd från detektorn, vilket försvårar tolkningen av signalen. För FTIR förväntas ytan inte av lika stor betydelse eftersom ljuset passerar rakt igenom mediet. Raman-modellerna presterade väldigt olika för olika kemiska komponenter och mellan de två olika provmaterialen, pulver och borrhärnor (Tabell 6). För träpulverproverna gavs generellt väldigt höga  $R^2$ . Det var inget oväntat och stämde överens med hypotesen träpulverprover skulle prestera bättre. Dock gav borrhärnorna för Raman en god modell för hemicellulosa vilket stred mot hypotesen.

Modellen som gavs var däremot på gränsen till överanpassad. Modellerna som Raman med borrhärd gav för glukos och xylos var svaga. Även cellulosa modellen låg på gränsen till att se som svag. Extraktivämnenas modell för borrhärderna var däremot tydlig överanpassad.

En utmaning vid analys av Ramanspektrumet från borrhärderna var att välja en punkt hos veden som ansågs vara representativ. Vår- och sommarved har stora skillnader i kemisk sammansättning och kan därför ge väldigt olika resultat. Den kemiska analysen baserades på hela regioner av borrhärderna, inklusive vår och sommarved. Därför är det intressant att veta hur en godtycklig punkt inom denna region kan användas för att skapa en modell. Träpulver från en blandning av vår- och sommarved förväntades ge bättre resultat än en slumpvis utvald punkt hos den råa borrhärden. Nackdelen med träpulvret ligger istället hos förarbetet som krävs för att ta fram det. Om träpulvret verkligen gav en signifikant bättre representation av trädets vedkemi än borrhärderna förväntas dess modeller generellt ha ett högre  $R^2$ . De enda modeller som direkt kunde jämföras var för glukos, cellulosa och hemicellulosa (Tabell 6). Som förväntat presterade träpulverproverna bättre för alla tre av dessa ämnen. Skillnaden mellan träpulver- och borrhärdproverna var särskilt stor för glukos, med  $R^2$  på 93 % mot 27 %. För cellulosa och hemicellulosa var skillnaden inte lika signifikant. Dessa två makromolekyler anses vara centrala vid förutsägning av vedkemi, eftersom de tillsammans utgör en stor andel av den sekundära cellväggen (Zhong & Ye 2015). Att borrhärderna endast presterade lite sämre än träpulverproverna tyder på deras användbarhet som framtida provmaterial. Visuellt framstod de båda förbehandlade spektrumet som relativt lika (Figur 5). Den enda stora skillnaden tycks ligga i intervallet 1200-1400  $\text{cm}^{-1}$ .

Modellerna i denna studie skapades utifrån spektraldata från 15 olika träprover. Dessa 15 prover kommer dock från endast 5 olika träd. Tidigare studier med FTIR har framgångsrikt skapat modeller för ligninhalten hos ved med prover från 44 olika granar, från tre olika lokaler (Raïskila et al. 2007). Gifty et al. (2016) nämner i en liknande studie att dess provstorlek på 40 prover ansågs vara relativt litet. Som tidigare nämnt är korsvalidering starkt beroende av provstorleken. Att endast 5 träd användes medför att det kan ifrågasättas hur representativ studien är.

En pålitlig modell för vedkemi bör vara kalibrerad på ett sätt som kan hantera den naturliga variationen hos olika träd. Detta innebär bland annat att modellens styrka är beroende av hos de ingående provernas räckvidd, det vill säga deras max- och minimivärde för de kemiska komponenterna, eftersom dessa avgör inom vilka ramar som modellen bäst kan appliceras. Med en liten provstorlek medföljer också att räckvidden riskerar att bli lägre. Li et al. (2015) använde en provstorlek på totalt 171 prover för att med närinfraröd spektroskopi skapa en modell för cellulosa, hemicellulosa och lignin hos bambu med goda resultat. Variationen hos proverna var: för cellulosa 37,29-53,7 % med en SD på 2,90, för hemicellulosa 17,70-23,65 % med en SD på 2,53, för lignin 13,82-23,86 % med en SD på 1,84. I denna studie var variationen (Tabell 2) hos proverna: för cellulosa 35,39-44,39 % med en SD på 2,37, för hemicellulosa 19,40-23,65 % med en SD på 1,22, och för lignin 26,60-31,40 % med en SD på 1,58. Variationen var alltså betydligt lägre hos proverna i denna studie. Monosackariderna uppvisar en ännu lägre variation där SD inte ens överstiger 1, med undantag för glukos.

Provstorleken påverkar också vilken signifikans som tilldelas avvikande värden. Ett sådant potentiellt avvikande prov framgår i Figur 4. Där framgår att formen hos ett spektrum, färgat ljusblått, ofta avviker från det övriga. På grund av studiens låga provstorlek kommer detta spektrum att kraftigt påverka provernas totala variation. Att ta bort avvikande prover skulle dock begränsa vilka värden som modellen kan förutsäga. I denna studie skulle det även minska provstorleken med hela 6,7 % per avvikelse. Vad som kan anses vara ett avvikande

värde är inte heller självklart. Ett avvikande värde bland 15 prover skulle exempelvis kunna vara fullt normalt bland 100 prover. Även om borttagning av ett eller två avvikande värden i vissa fall kunde öka modellens förklaringsgrad kan det alltså ifrågasättas ifall detta verkligen var motiverat.

Även den kemiska analysen påverkades av den lilla provstorleken, genom att flera modifikationer fick göras. De exakta protokoll som följdes vid den kemiska analysen av MoRe är inte kända. Hur detta kan ha påverkat kvalitén hos analysen är därför svårt att uppskatta. Utifrån procenthalten av cellulosa, hemicellulosa och lignin (Tabell 1 & 2) tycktes dock inga iögonfallande fel ha uppstått, utan värdena (35,39 – 44,39 % cellulosa, 19,40 – 23,65 % hemicellulosa och 26,60–31,40 % lignin) faller väl inom ramen för vad som kan förväntas av granens vedkemi (Raiskila et al. 2007).

Generellt eftersträvas att så få faktorer som möjligt används för modellen, vilket är varför någon sorts krav på signifikans ofta är nödvändigt. Antalet faktorer hos modellerna i denna studie var dock maximalt 5. Jämförelsevis gavs lägsta PRESS-värde för Raiskila et al. (2007) vid maximalt 16 faktorer vid modellering av lignin hos grankloner, varav de modeller med fler än 7-8 faktorer valdes bort för att undvika överanpassning. Även då alla faktorer extraherades i denna studie blev modellerna sällan tydligt överanpassade. Därför ansågs det vara motiverat att ta bort kravet för pval hos modellerna och använda alla tillgängliga faktorer.

Även om modellerna som skapades i denna studie inte var till någon realistisk kommersiell nytta visade ändå resultatet på teknikens potential. Både Raman och FTIR lyckades reproducera flera spektrum med liknande karaktär (Figur 4 & 5). Eftersom ingen av metoderna klarade av att skapa fullständiga modeller för alla kemiska komponenter var det svårt att göra en tydlig och rättvis jämförelse mellan dem. Baserat på tillgängliga  $R^2$  gav FTIR bäst modeller för mannos och lignin, och Raman (både träpulver och borkärnor) för arabinos, glukos, xylos, extraktivämnen, cellulosa och hemicellulosa. Ingen av metoderna lyckades inkludera någon faktor för galaktos vid lägsta PRESS-värde. Då antalet faktorer som inkluderades vid lägsta PRESS-värde var noll kom metoderna fram till exakt samma PRESS-värde för samma kemiska komponent (Tabell 3 & 5). Det är dock svårt att utläsa någon värdefull information från detta eftersom PRESS-värde används som ett mått för att välja ut ett optimalt antal faktorer.

## Slutsatser

Varken FTIR eller Raman lyckades skapa en modell för varje kemisk komponent. De modeller som skapades hade ett lågt antal faktorer, med varierande förklaringsgrad. Metoderna lyckades däremot skapa modeller för några kemiska komponenter. FTIR lyckades skapa modeller med hög förklaringsgrad för mannos och lignin (Tabell 4), och Raman för arabinos, glukos, cellulosa och hemicellulosa (Tabell 6). Således tycks metoderna kunna komplettera varandra, dock inte som förväntat. Val av normalisering visade sig kunna ha en betydande effekt på modellens förklaringsgrad för enskilda kemiska komponenter. Inom Raman ansågs träpulverproverna prestera lite bättre som testmaterial än borkärnesproverna, i enlighet med hypotesen. Dock framgick det inte hur stor skillnaden var i förhållande till det extra arbete som krävs för att mala fram träpulver. Svagheten hos studien tillskrevs framförallt den låga provstorleken, vilken påverkade den kemiska analysen samt ifall de ingående proverna kan tolkas som representativa. Både FTIR- och Ramanspektroskopi anses fortfarande vara av intresse för vidare studier med större provstorlekar.

## REFERENSER

The Penguin Group. (1977). The Penguin Dictionary of Physics. 2:a uppl. London. Penguin Books.

Agarwal, U.P. & Ralph, S.A. (1997). FT-Raman spectroscopy of wood: identifying contributions of lignin and carbohydrate polymers in the spectrum of black spruce (*Picea mariana*). *APPLIED SPECTROSCOPY*, volym 51(11), ss. 1648-1655. doi: 10.1366/0003702971939316.

Alcantara G, A., Honda, N, K., Ferreira, M, M, C. & Ferreira, A, G. (2007). Chemometric analysis applied in 1H HR-MAS NMR and FT-IR data for chemotaxonomic distinction of intact lichen samples. *Analytica Chimica Acta*, volym 595(2007), ss. 3–8. Doi:10.1016/j.aca.2007.03.032

Baratieri, M., Baggio, P., Fiori, L. & Grigante, M. (2008). Biomass as an energy source: Thermodynamic constraints on the performance of the conversion process. *Bioresource Technology*, volym 99, ss. 7063–7073. DOI: 10.1016/j.biortech.2008.01.006.

Binder J.B. & Raines R.T. (2009). Simple chemical transformation of lignocellulosic biomass into furans for fuels and chemicals. *Journal of the American Chemical Society*, volym 131, ss. 1979–1985.

Chen, H., Ferrari, C. & Bramanti, E. (2010). Qualitative and quantitative analysis of wood samples by Fourier transform infrared spectroscopy and multivariate analysis. *Carbohydrate Polymers*, volym 82(2010), ss. 772-778. Doi: 10.1016/j.jcarbpol.2010.05.052.

Ellingwood, B. R., Rosowsky, D. V., Li, Y. & Kim, J. H. (2004). Fragility assessment of light-frame wood construction subjected to wind and earthquake hazards. *Journal of Structural Engineering*, volym 130(12), ss.1921–1930. DOI: 10.1061/(ASCE)0733-9445(2004)130:12(1921).

Felten, J., Hall, H., Jaumot, J., Tauler, R., Juan, A.D. & Gorzsás., A. (2015). Vibrational spectroscopic image analysis of biological material using multivariate curve resolution–alternating least squares (MCR-ALS). *Nature Protocols*, volym 10, ss. 217–240. DOI: 10.1038/nprot.2015.008.

Fridman, J., Stendahl, J., Nilsson P., H. & Ståhl G. (2017). *Forest statistics 2017, Official Statistics of Sweden*. Swedish University of Agricultural Sciences. Umeå, Johan Fransson. ISSN 0280-0543.

Ghaffar, S.H. & Fan, M. (2013). Structural analysis for lignin characteristics in biomass straw. *Biomass and Bioenergy*, volym 57, ss. 264-279. DOI: 10.1016/j.biombioe.2013.07.015.



- Gierlinger, N. (2014). Revealing changes in molecular composition of plant cell walls on the micron-level by Raman mapping and vertex component analysis (VCA). *FRONTIERS IN PLANT SCIENCE*, volym 5, Article 306, ss. 1-10. DOI: 10.3389/fpls.2014.00306.
- Gifty E. Acquah, Brian K. Via, Oladiran O. Fasina, & Lori G. Eckhardt. (2016). Rapid Quantitative Analysis of Forest Biomass Using Fourier Transform Infrared Spectroscopy and Partial Least Squares Regression. *Journal of Analytical Methods in Chemistry*, volym 2016, Article ID 1839598, ss. 1-10. DOI:10.1155/2016/1839598.
- Gorzás, A. (2018). Vibrational Spectroscopy – User Licence Course – 2 ECTS, *grundkurs i vibrationsspektroskopi*, Umeå Universitet.
- Hames, B.R, Thomas S.R, Sluiter A.D, Roth C.J. & Templeton D.W. (2003). Rapid biomass analysis- New tools for compositional analysis of corn stover feedstocks and process intermediates from ethanol production. *Applied biochemistry and biotechnology*, volym 105, ss. 5–16. DOI: 10.1385/ABAB: 105:1-3:5.
- Jiang, W., Han, G., Zhang, Y. & Wang, M. (2010). Fast compositional analysis of ramie using near-infrared spectroscopy. *Carbohydrate Polymers*, volym 81(4), ss. 937-941. DOI: 10.1016/j.carbpol.2010.04.009.a
- Kumar, M., Campbell, L. & Turner, S. (2016). Secondary cell walls: biosynthesis and manipulation. *Journal of Experimental Botany*, volym 67(2), ss. 515–531. DOI: 10.1093/jxb/erv533.
- Li, X., Sun, C., Zhou, B. & He, Y. (2015). Determination of Hemicellulose, Cellulose and Lignin in Moso Bamboo by Near Infrared Spectroscopy. *Scientific reports*, volym 5(17210), ss. 1-11. DOI: 10.1038/srep17210.
- Lopez-Reyes, G., & Rull Pérez, F. (2017). A method for the automated Raman spectra acquisition. *J. Raman Spectroscopy*, volym 48, ss. 1654–1664. DOI: 10.1002/jrs.5185.
- Lupoi, J.S. (2012). Developments in enzyme immobilization and near-infrared Raman spectroscopy with downstream renewable energy applications. Graduate Theses and Dissertations. Iowa State University. Paper 12732.
- Lupoi J.S. Gjersing E. & Davis M.F. (2015). Evaluating Lignocellulosic Biomass, Its Derivatives, and Downstream Products with Raman Spectroscopy. *Frontiers in Bioengineering and Biotechnology*, volym 3(50), ss. 1-18. DOI: 10.3389/fbioe.2015.00050.
- Lupoi, J.S., Singh, S., Parthasarathi, R., Simmons, B.A. & Henry, R.J. (2015). Recent innovations in analytical methods for the qualitative and quantitative assessment of lignin. *Renewable and Sustainable Energy Reviews*, volym 49, ss. 871-906. DOI: 10.1016/j.rser.2015.04.091.

Mannan, H. R. & McNeil, J. J. (2013). Computer programs to estimate overoptimism in measures of discrimination for predicting the risk of cardiovascular diseases. *Journal of Evaluation in Clinical Practice*, volym 19, ss. 358-362. DOI: 10.1111/j.1365-2753.2012.01834.x.

Raiskila, S., Pulkkinen, M., Laakso, T., Fagerstedt, K., Löijä, M., Mahlberg, R., Paajanen, L., Ritschkoff, A.-C. & Saranpää, P. (2007). FTIR spectroscopic prediction of Klason and acid soluble lignin variation in Norway spruce cutting clones. *Silva Fennica*, volym 41(2), ss. 351–371. DOI: 10.14214/sf.301.

Ramage, M.H., Burridge, H., Busse-Wicher, M., Fereday, G., Reynolds, T., Shah, D.U., Wu, G., Yu, L., Fleming, P., Densley-Tingley, D., Allwood, J., Dupree, P., Linden, P.F. & Scherman, O. (2017). The wood from the trees: The use of timber in construction. *Renewable and Sustainable Energy Reviews*, volym 68, ss. 333-359. DOI:10.1016/j.rser.2016.09.107.

Rojo, E., Alonso, M., Domínguez, J., Saz-Orozco, B., Oliet, M. & Rodriguez, F. (2013). Alkali treatment of viscose cellulosic fibers from eucalyptus wood: Structural, morphological, and thermal analysis. *Journal of Applied Polymer Science*, volym 130(3), ss. 2198-2204. DOI: 10.1002/app.39399.

Roggo, Y., Chalus, P., Maurer, L., Lema-Martinez, C., Edmond, A. & Jent, N. (2007). A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies. *Journal of Pharmaceutical and Biomedical Analysis*, volym 44(3), ss. 683-700. DOI: 10.1016/j.jpba.2007.03.023.

Runcorn, T.H., Görlitz, F.G., Murray, R.T. & Kelleher, E.J.R. (2018). Visible Raman-Shifted Fiber Lasers for Biophotonic Applications. *Journal of Selected Topics in Quantum Electronics*, volym. 24(3), ss. 1-8. DOI: 10.1109/JSTQE.2017.2770101.

SAS Institute Inc. (2008). The PLS Procedure I. *SAS/STAT® 9.2 User's Guide*. Cary, NC: SAS Institute Inc, USA, ss. 4759-4805.

Schwanninger, M., Rodrigues, J.C. & Fackler, K. (2011). A Review of Band Assignments in near Infrared Spectra of Wood and Wood Components. *Journal of Near Infrared Spectroscopy*. volym 19(5), ss. 287-308. DOI: 10.1255/jnirs.955.

Séne, C, F, B., McCann M, C., Wilson, R, H. & Grinter R. (1994). Fourier-transform Raman and Fourier-transform infrared spectroscopy (an investigation of five higher plant cell walls and their components). *Plant Physiol*, volym 1994(106), ss.1623–31. doi: org/10.1104/pp.106.4.1623.

SWINEHART, D. F. (1962). The beer-lambert law. *Journal of chemical education*, volym 39(7), ss. 333-335. DOI: 10.1021/ed039p333.

- Weinstock, I. A., Atalla, R. H., Reiner, R. S., Moen, M. A., Hammel, K. E., Houtman, C. J., et al. (1997). A new environmentally benign technology for transforming wood pulp into paper - Engineering polyoxometalates as catalysts for multiple processes. *Journal of Molecular Catalysis A: Chemical*, volym 116(1), ss. 59–84. DOI: 10.1016/S1381-1169(96)00074-X.
- Xu, F., Yu, J., Tesso, T., Floyd-Dowell, F. & Wang, D. (2013). Qualitative and quantitative analysis of lignocellulosic biomass using infrared techniques: A mini-review. *Applied Energy*, volym 104, ss. 801-809. DOI: 10.1016/j.apenergy.2012.12.019.
- Zeng, Y., Yarbrough, J.M., Mittal, A. et al. (2016). In situ label-free imaging of hemicellulose in plant cell walls using stimulated Raman scattering microscopy. *Biotechnology for Biofuels*, volym 9(256), ss. 1-16. DOI: 10.1186/s13068-016-0669-9.
- Zhong, R. & Ye, Z.H. (2015). Secondary Cell Walls: Biosynthesis, Patterned Deposition and Transcriptional Regulation. *Plant and Cell Physiology*, volym 56(2), ss. 195–214. DOI: 10.1093/pcp/pcu140.
- Zhang, H., Wang, X. & Li, Y. (2018). Measuring radiative properties of silica aerogel composite from FTIR transmittance test using KBr as diluents. *Experimental Thermal and Fluid Science*, volym 91, ss. 144-154. DOI: 10.1016/j.expthermflusci.2017.10.010
- Zhou, C., Jiang, W., Via, B.K., Fasina, O. & Han, G. (2015). Prediction of mixed hardwood lignin and carbohydrate content using ATR-FTIR and FT-NIR. *Carbohydrate Polymers*, volym 121, ss. 336-341. DOI: 10.1016/j.carbpol.2014.11.062.